Optimization for Quantum Information Science Problems

Jeffrey Larson

Argonne National Laboratory

March 3, 2021

Optimization in Quantum Information Systems

- My background is solving difficult numerical optimization problems arising the quantum space.

Optimization in Quantum Information Systems

- My background is solving difficult numerical optimization problems arising the quantum space.

- I'll talk about four today:

Optimization in Quantum Information Systems

- My background is solving difficult numerical optimization problems arising the quantum space.

- I'll talk about four today:
  - Maximizing concurrence with Otten (HRL), Pelton (UMBC), Min, Wild, and Gray (ANL)

Optimization in Quantum Information Systems

- My background is solving difficult numerical optimization problems arising the quantum space.

- I'll talk about four today:
  - Maximizing concurrence with Otten (HRL), Pelton (UMBC), Min, Wild, and Gray (ANL)
  - Optimal Circuit Cutting with Tang, Tomesh, Martonosi (Princeton), and Suchara (ANL)

Optimization in Quantum Information Systems

- My background is solving difficult numerical optimization problems arising the quantum space.

- I'll talk about four today:
  - Maximizing concurrence with Otten (HRL), Pelton (UMBC), Min, Wild, and Gray (ANL)
  - Optimal Circuit Cutting with Tang, Tomesh, Martonosi (Princeton), and Suchara (ANL)
  - Design of a fixed frequency quantum processor with Morvan and Chen (LBNL)

Optimization in Quantum Information Systems

- My background is solving difficult numerical optimization problems arising the quantum space.

- I'll talk about four today:
  - Maximizing concurrence with Otten (HRL), Pelton (UMBC), Min, Wild, and Gray (ANL)
  - Optimal Circuit Cutting with Tang, Tomesh, Martonosi (Princeton), and Suchara (ANL)
  - Design of a fixed frequency quantum processor with Morvan and Chen (LBNL)
  - ? Maximizing quantum Fisher information

Outline

Maximizing concurrence

- Entanglement is delicate thing and can be easily destroyed

Maximizing concurrence

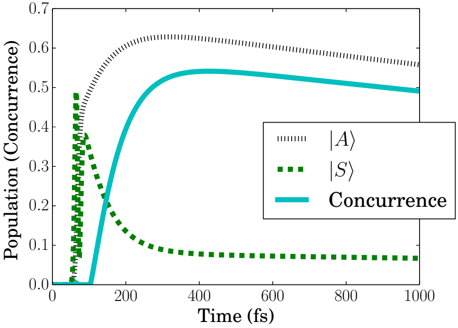
- Entanglement is delicate thing and can be easily destroyed
- Concurrence is a measure of entanglement of a quantum system

Maximizing concurrence

- Entanglement is delicate thing and can be easily destroyed
- Concurrence is a measure of entanglement of a quantum system
- Concurrence of two quantum dots excited by a single optical laser pulse:

Maximizing concurrence

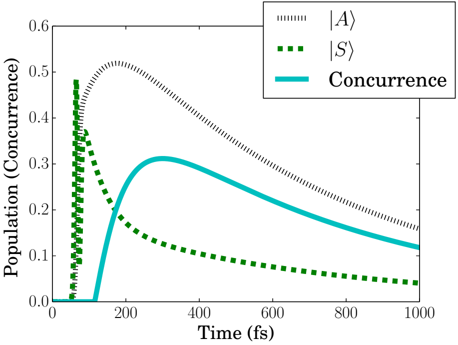
- Entanglement is delicate thing and can be easily destroyed
- Concurrence is a measure of entanglement of a quantum system
- Concurrence of two quantum dots excited by a single optical laser pulse:



"QuaC: Parallel time Dependent Open Quantum Systems Solver." Otten

Maximizing concurrence

- Entanglement is delicate thing and can be easily destroyed
- Concurrence is a measure of entanglement of a quantum system
- Concurrence of two quantum dots excited by a single optical laser pulse:



"QuaC: Parallel time Dependent Open Quantum Systems Solver." Otten

Maximizing concurrence

- Pairwise concurrence is measured by

$$C_{ij} = \max\{0, \sqrt{\lambda_1} - \sqrt{\lambda_2} - \sqrt{\lambda_3} - \sqrt{\lambda_4}\},$$

  where λ_k are the (descending) eigenvalues of a density matrix relating particles i and j.

Maximizing concurrence

- Pairwise concurrence is measured by

$$C_{ij} = \max\{0, \sqrt{\lambda_1} - \sqrt{\lambda_2} - \sqrt{\lambda_3} - \sqrt{\lambda_4}\},$$

  where λ_k are the (descending) eigenvalues of a density matrix relating particles i and j.
- Possible goal: Identify quantum system parameters x solving

Maximizing concurrence

- Pairwise concurrence is measured by

$$C_{ij} = \max\{0, \sqrt{\lambda_1} - \sqrt{\lambda_2} - \sqrt{\lambda_3} - \sqrt{\lambda_4}\},$$

  where λ_k are the (descending) eigenvalues of a density matrix relating particles i and j.

- Possible goal: Identify quantum system parameters x solving

Maximizing concurrence

- Pairwise concurrence is measured by

$$C_{ij} = \max\{0, \sqrt{\lambda_1} - \sqrt{\lambda_2} - \sqrt{\lambda_3} - \sqrt{\lambda_4}\},$$

  where λ_k are the (descending) eigenvalues of a density matrix relating particles i and j.

- Possible goal: Identify quantum system parameters x solving

$$\underset{x}{\text{maximize}} \sum_{ij} C_{ij}(x)^2$$

Maximizing concurrence

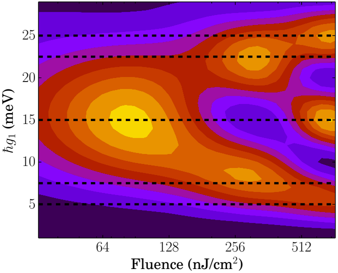
- Pairwise concurrence is measured by

$$C_{ij} = \max\{0, \sqrt{\lambda_1} - \sqrt{\lambda_2} - \sqrt{\lambda_3} - \sqrt{\lambda_4}\},$$

  where λ_k are the (descending) eigenvalues of a density matrix relating particles i and j.

- Possible goal: Identify quantum system parameters x solving

$$\underset{x}{\text{maximize}} \sum_{ij} C_{ij}(x)^2$$



Otten, Larson, Min, Wild, Pelton, Gray. Origins and optimization of entanglement in plasmonically coupled quantum dots. Physical Review A, 2016

Possible extensions

- Concurrence is not computably defined for odd numbers of qubits (there are formulas, but they are hard to compute numerically).

Possible extensions

- Concurrence is not computably defined for odd numbers of qubits (there are formulas, but they are hard to compute numerically).
- There is a computable definition for even numbers of qubits which could be used.

Possible extensions

- Concurrence is not computably defined for odd numbers of qubits (there are formulas, but they are hard to compute numerically).
- There is a computable definition for even numbers of qubits which could be used.
- Higher dimensional entanglement is less well understood theoretically, but there are some special states would be interesting to try and create.

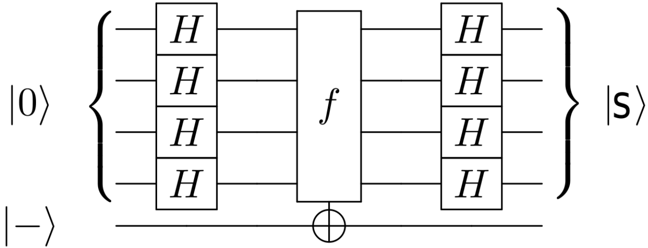
Outline

Real-world circuit fidelities

Take an n-qubits quantum computer from IBM and run Bernstein-Vazirani algorithm using $\lfloor \frac{n}{2} \rfloor$ qubits.

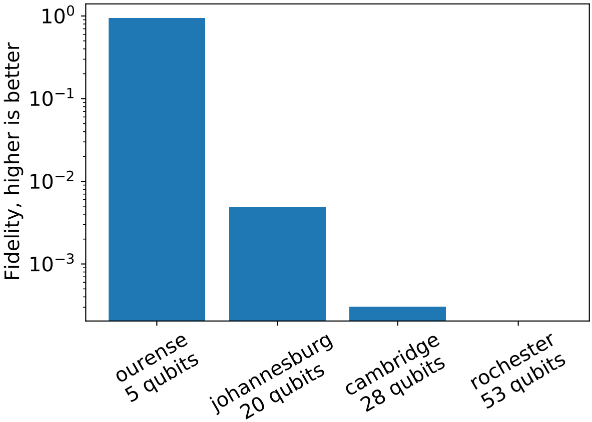
Real-world circuit fidelities

Take an n-qubits quantum computer from IBM and run Bernstein-Vazirani algorithm using $\lfloor \frac{n}{2} \rfloor$ qubits.



Real-world circuit fidelities

Take an n-qubits quantum computer from IBM and run Bernstein-Vazirani algorithm using $\lfloor \frac{n}{2} \rfloor$ qubits.



Can we help this?

- The set of quantum circuits that can be reliably run on NISQ devices is limited by their noisy operations and low qubit counts.

Can we help this?

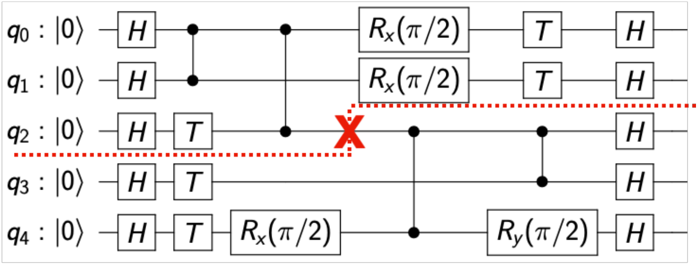
- The set of quantum circuits that can be reliably run on NISQ devices is limited by their noisy operations and low qubit counts.
- We consider a hybrid classical/quantum computing approach

Can we help this?

- The set of quantum circuits that can be reliably run on NISQ devices is limited by their noisy operations and low qubit counts.
- We consider a hybrid classical/quantum computing approach
- Cuts large quantum circuits into smaller subcircuits

Can we help this?

- The set of quantum circuits that can be reliably run on NISQ devices is limited by their noisy operations and low qubit counts.
- We consider a hybrid classical/quantum computing approach
- Cuts large quantum circuits into smaller subcircuits
- Classical post-processing can then reconstruct the output of the original circuit.



Modeling variables

$$y_{v,c} \equiv \begin{cases} 1 & \text{if vertex } v \text{ is in subcircuit } c \\ 0 & \text{otherwise} \end{cases}, \ \forall v \in V, \forall c \in C$$

Modeling variables

$$y_{v,c} \equiv \begin{cases} 1 & \text{if vertex } v \text{ is in subcircuit } c \\ 0 & \text{otherwise} \end{cases}, \forall v \in V, \forall c \in C$$

$$x_{e,c} \equiv \begin{cases} 1 & \text{if edge } e \text{ is cut by subcircuit } c \\ 0 & \text{otherwise} \end{cases}, \forall e \in E, \forall c \in C$$

Modeling variables

$$y_{v,c} \equiv \begin{cases} 1 & \text{if vertex } v \text{ is in subcircuit } c \\ 0 & \text{otherwise} \end{cases}, \forall v \in V, \forall c \in C$$

$$x_{e,c} \equiv \begin{cases} 1 & \text{if edge } e \text{ is cut by subcircuit } c \\ 0 & \text{otherwise} \end{cases}, \forall e \in E, \forall c \in C$$

The number of qubits required to run a subcircuit is the sum of:

- The number of original input qubits
- The number of initialization qubits induced by cutting

Modeling variables

- The number of original input qubits is $\alpha_c \equiv \sum_{v \in V} w_v \times y_{v,c}, \forall c \in C$, where $w_v \in \{0, 1, 2\}$ is the number of original input qubits directly connected to $v \in V$.

Modeling variables

- The number of original input qubits is $\alpha_c \equiv \sum_{v \in V} w_v \times y_{v,c}, \forall c \in C$, where $w_v \in \{0, 1, 2\}$ is the number of original input qubits directly connected to $v \in V$.

- The number of initialization qubits is: $\rho_c \equiv \sum_{e:(e_a,e_b) \in E} x_{e,c} \times y_{e_b,c}, \forall c \in C$.

Modeling variables

- The number of original input qubits is $\alpha_c \equiv \sum_{v \in V} w_v \times y_{v,c}, \forall c \in C$, where $w_v \in \{0, 1, 2\}$ is the number of original input qubits directly connected to $v \in V$.

- The number of initialization qubits is: $\rho_c \equiv \sum_{e:(e_a,e_b) \in E} x_{e,c} \times y_{e_b,c}, \forall c \in C$.

- The number of measurement qubits is

$$O_c \equiv \sum_{e:(e_a,e_b) \in E} x_{e,c} \times y_{e_a,c}, \forall c \in C.$$

Modeling variables

- The number of original input qubits is $\alpha_c \equiv \sum_{v \in V} w_v \times y_{v,c}, \forall c \in C$, where $w_v \in \{0, 1, 2\}$ is the number of original input qubits directly connected to $v \in V$.

- The number of initialization qubits is: $\rho_c \equiv \sum_{e:(e_a,e_b) \in E} x_{e,c} \times y_{e_b,c}, \forall c \in C$.

- The number of measurement qubits is

$$O_c \equiv \sum_{e:(e_a,e_b) \in E} x_{e,c} \times y_{e_a,c}, \forall c \in C.$$

Consequently, the number of qubits in a subcircuit that contributes to the final measurement of the original uncut circuit is

$$f_c \equiv \alpha_c + \rho_c - O_c, \forall c \in C.$$

Constraints

- Every vertex must be in a subcircuit

Constraints

- Every vertex must be in a subcircuit
- The required qubits for each circuit is limited.

Constraints

- Every vertex must be in a subcircuit
- The required qubits for each circuit is limited.
- Some symmetry-breaking constraints

Objective

- The number of cuts made is

$$K = \frac{1}{2} \sum_{c \in C} \sum_{e \in E} x_{e,c},$$

Objective

- The number of cuts made is

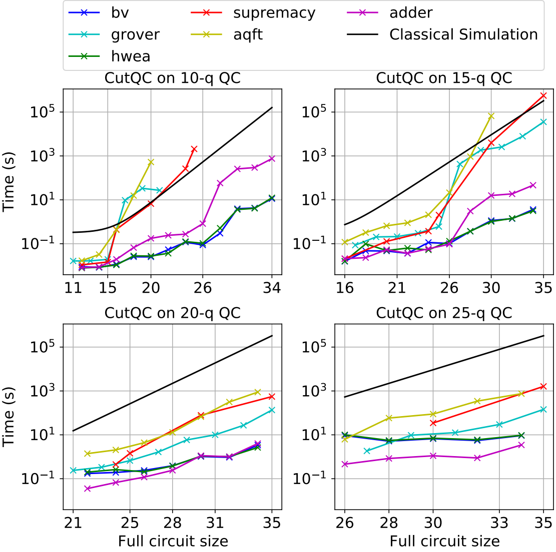
$$K = \frac{1}{2} \sum_{c \in C} \sum_{e \in E} x_{e,c},$$

- The objective function for the MIP cut searcher is reconstruction time estimator

$$L \equiv 4^K \sum_{c=2}^{n_C} \prod_{i=1}^{c} 2^{f_i},$$

which captures cost of building the full 2^n probabilities for a n-qubit uncut circuit

Results



Tang et al. "CutQC: Using Small Quantum Computers for Large Quantum Circuit Evaluations." ASPLOS21

Outline

Scaling up quantum devices is a challenge

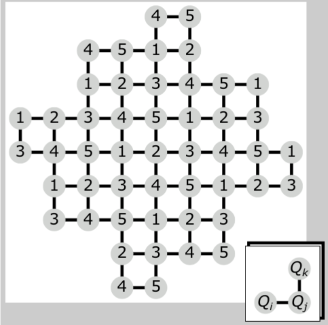
- Fixed-frequency transmons are an appealing technology due to their long coherence times ($\sim 100~\mu$s)

Scaling up quantum devices is a challenge

- Fixed-frequency transmons are an appealing technology due to their long coherence times ($\sim 100\ \mu$s)
- Scaling fixed-frequency architectures requires precise relative frequency requirements.

Scaling up quantum devices is a challenge

- Fixed-frequency transmons are an appealing technology due to their long coherence times ($\sim$100 μs)
- Scaling fixed-frequency architectures requires precise relative frequency requirements.
- Want to avoid collisions in frequencies.



Hertzberg et al., https://arxiv.org/pdf/2009.00781.pdf

Problem description

- A processor is "collision-free" when the various types of frequency collisions are avoided (with some margins) determined through modeling or experimentally

Problem description

- A processor is “collision-free” when the various types of frequency collisions are avoided (with some margins) determined through modeling or experimentally
- The yield measure the number of a potentially valid quantum processor

Problem description

- A processor is "collision-free" when the various types of frequency collisions are avoided (with some margins) determined through modeling or experimentally
- The yield measure the number of a potentially valid quantum processor
- Chips are fabricated in batches, and they want to have at least one valid chip per batch.

Frequency collisions can take a variety of forms

- f_i avoid the $0 \mapsto 1$ transitions of j:

$$|f_i - f_j| \geq \delta_1 \qquad \forall (i, j) \in E$$

Frequency collisions can take a variety of forms

- f_i avoid the $0 \mapsto 1$ transitions of j:

$$|f_i - f_j| \geq \delta_1 \qquad \forall (i,j) \in E$$

- Other types of collisions are more subtle: Can happen when the sum of frequency of the target and a neighbors is equal to the frequency of the $|0\rangle \to |2\rangle$ 2 photon transition.

$$|2f_i + \alpha_i - f_k - f_j| \geq \delta_7$$
$$\forall j,k \in N \; s.t. \; \exists i \in N \text{ with } (i,j) \in \vec{E} \text{ and } (i,k) \in \vec{E} \text{ or } (k,i) \in \vec{E}$$

Frequency collisions can take a variety of forms

- f_i avoid the $0 \mapsto 1$ transitions of j:

$$|f_i - f_j| \geq \delta_1 \qquad \forall (i,j) \in E$$

- Other types of collisions are more subtle: Can happen when the sum of frequency of the target and a neighbors is equal to the frequency of the $|0\rangle \rightarrow |2\rangle$ 2 photon transition.

$$|2f_i + \alpha_i - f_k - f_j| \geq \delta_7$$
$$\forall j,k \in N \ s.t. \ \exists i \in N \text{ with } (i,j) \in \vec{E} \text{ and } (i,k) \in \vec{E} \text{ or } (k,i) \in \vec{E}$$

- A possible objective:

$$\text{maximize} \sum_i w_i \delta_i$$

Frequency collisions can take a variety of forms

- f_i avoid the $0 \mapsto 1$ transitions of j:

$$|f_i - f_j| \geq \delta_1 \qquad \forall (i,j) \in E$$

- Other types of collisions are more subtle: Can happen when the sum of frequency of the target and a neighbors is equal to the frequency of the $|0\rangle \to |2\rangle$ 2 photon transition.

$$|2f_i + \alpha_i - f_k - f_j| \geq \delta_7$$
$$\forall j,k \in N \ s.t. \ \exists i \in N \text{ with } (i,j) \in \vec{E} \text{ and } (i,k) \in \vec{E} \text{ or } (k,i) \in \vec{E}$$

- A possible objective:

$$\text{maximize} \sum_i w_i \delta_i$$

Frequency collisions can take a variety of forms

- f_i avoid the $0 \mapsto 1$ transitions of j:

$$|f_i - f_j| \geq \delta_1 \qquad \forall (i,j) \in E$$

- Other types of collisions are more subtle: Can happen when the sum of frequency of the target and a neighbors is equal to the frequency of the $|0\rangle \rightarrow |2\rangle$ 2 photon transition.

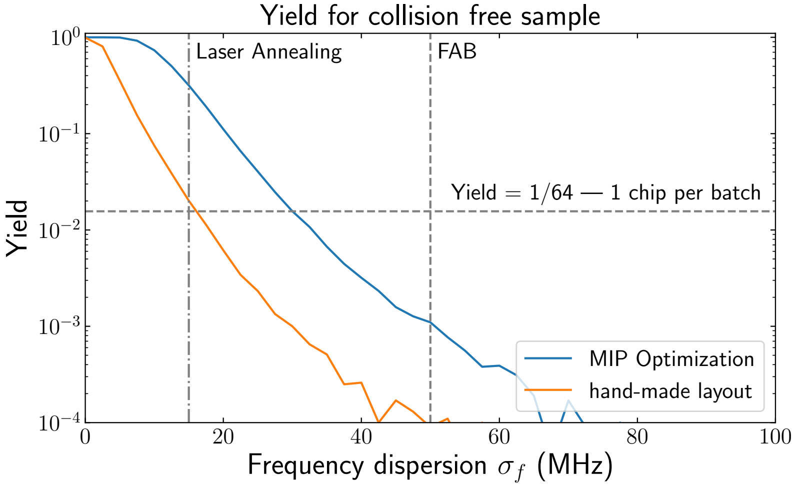
$$|2f_i + \alpha_i - f_k - f_j| \geq \delta_7$$
$$\forall j,k \in N \ s.t. \ \exists i \in N \text{ with } (i,j) \in \vec{E} \text{ and } (i,k) \in \vec{E} \text{ or } (k,i) \in \vec{E}$$

- A possible objective:

$$\text{maximize} \sum_i w_i \delta_i$$

with $\delta_i \geq \bar{\delta}_i$.

Two solutions on 6-node ring



Active extensions

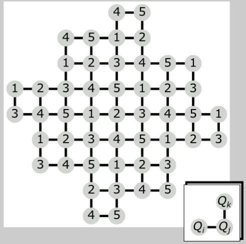
- Accounting for qutrits

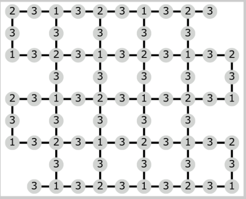
Active extensions

- Accounting for qutrits
- Identifying an optimal chunk that allows for massive designs

Active extensions

- Accounting for qutrits
- Identifying an optimal chunk that allows for massive designs
- Trying to optimize the connectivity in the graph. (For now, just assigning frequencies to a given architecture.)





Outline

Maximizing Concurrence

Optimal Circuit Cutting

Fixed-frequency quantum processor

Maximizing Quantum Fisher Information

Quantum Fisher information

- A quantum analogue to classical Fisher information (which describes how sensitive a model is changes in a parameter.)

Quantum Fisher information

- A quantum analogue to classical Fisher information (which describes how sensitive a model is changes in a parameter.)

- A central quantity in quantum sensing.

Quantum Fisher information

- A quantum analogue to classical Fisher information (which describes how sensitive a model is changes in a parameter.)

- A central quantity in quantum sensing.
  - Classical: for n independent sensors sensing a physical parameter, the precision is improved $O(\frac{1}{\sqrt{n}})$.

Quantum Fisher information

- A quantum analogue to classical Fisher information (which describes how sensitive a model is changes in a parameter.)

- A central quantity in quantum sensing.
  - Classical: for n independent sensors sensing a physical parameter, the precision is improved $O(\frac{1}{\sqrt{n}})$.
  - Quantum: the precision is improved $O(\frac{1}{n})$.

Quantum Fisher information

- A quantum analogue to classical Fisher information (which describes how sensitive a model is changes in a parameter.)

- A central quantity in quantum sensing.
  - Classical: for n independent sensors sensing a physical parameter, the precision is improved $O(\frac{1}{\sqrt{n}})$.
  - Quantum: the precision is improved $O(\frac{1}{n})$.

- For a state with density matrix $\rho(x) = \sum_i^N \lambda_i |\psi_i\rangle \langle\psi_j|$, the QFI is

$$\mathcal{F}(\rho(x), H) = \sum_{i,j} \frac{\lambda_i - \lambda_j}{2(\lambda_i + \lambda_j)} |\langle\psi_i| H |\psi_j\rangle|$$

Quantum Fisher information

- A quantum analogue to classical Fisher information (which describes how sensitive a model is changes in a parameter.)

- A central quantity in quantum sensing.
  - Classical: for n independent sensors sensing a physical parameter, the precision is improved $O(\frac{1}{\sqrt{n}})$.
  - Quantum: the precision is improved $O(\frac{1}{n})$.

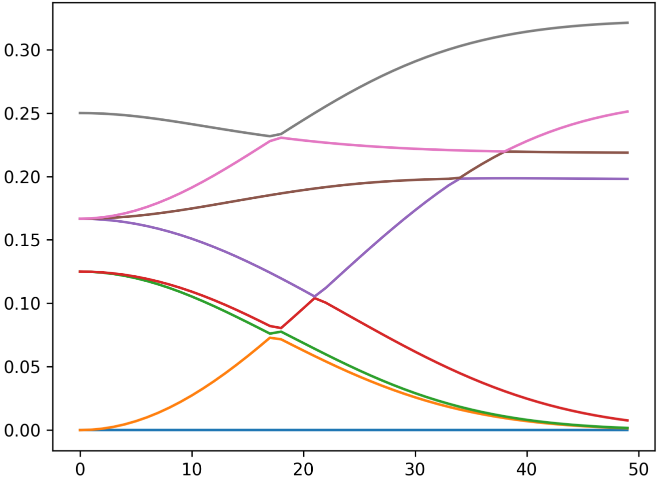
- For a state with density matrix $\rho(x) = \sum_i^N \lambda_i |\psi_i\rangle \langle\psi_j|$, the QFI is

$$\mathcal{F}(\rho(x), H) = \sum_{i,j} \frac{\lambda_i - \lambda_j}{2(\lambda_i + \lambda_j)} |\langle\psi_i| H |\psi_j\rangle|$$

- For large N, computing the QFI can be prohibitively difficult. Many papers maximize (upper) bounds of QFI

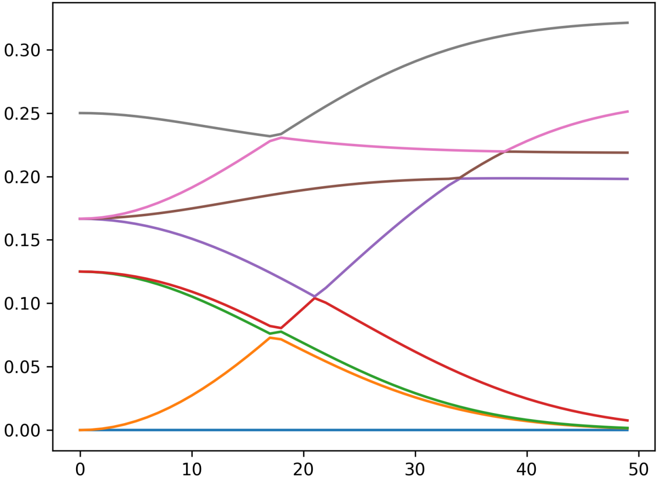
Mathematical fun

Example with $N = 8$. Take a starting point x_0 and a random direction d. Compute eigen-decomposition for $\rho(x_0 + \alpha_i d)$ and plot eigenvalues



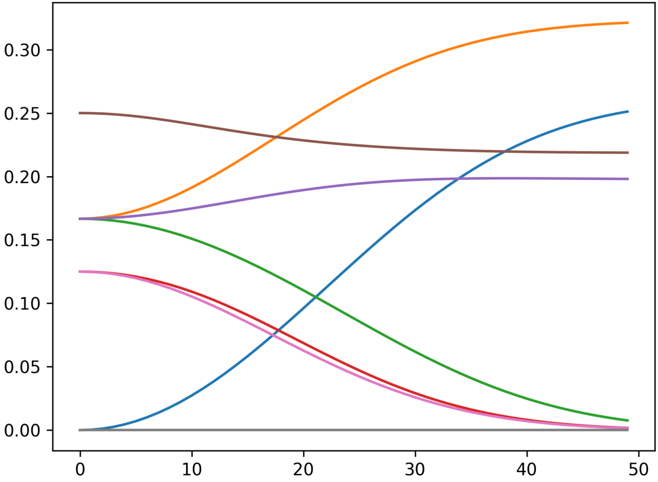
Mathematical fun

Instead, let's number of the eigenvalues of $\rho(x_0)$. Number the eigenvalues of $\rho(x_0 + \alpha_i d)$, using the eigenpairs at $\rho(x_0 + \alpha_{i-1} d)$



Mathematical fun

Instead, let's number of the eigenvalues of $\rho(x_0)$. Number the eigenvalues of $\rho(x_0 + \alpha_i d)$, using the eigenpairs at $\rho(x_0 + \alpha_{i-1} d)$



Further questions

- Can we get $\frac{d}{dx}\rho(x)$ for some parameters (possible for pulse parameters).

Further questions

- Can we get $\frac{d}{dx}\rho(x)$ for some parameters (possible for pulse parameters).

- For arbitrary H, find optimal states ρ and hope that their symmetries/generalizations "scale". (Can happen.) Show improved optimization on such H (and leave it to the physicist to find systems so that such results "scale".)

Further questions

- Can we get $\frac{d}{dx}\rho(x)$ for some parameters (possible for pulse parameters).
- For arbitrary H, find optimal states ρ and hope that their symmetries/generalizations "scale". (Can happen.) Show improved optimization on such H (and leave it to the physicist to find systems so that such results "scale".)
- Some H/ρ pairs may have analytic forms for their eigenpairs (but we really need cross-products to be "nice"). What could we do in that case?

Further questions

- Can we get $\frac{d}{dx}\rho(x)$ for some parameters (possible for pulse parameters).
- For arbitrary H, find optimal states ρ and hope that their symmetries/generalizations "scale". (Can happen.) Show improved optimization on such H (and leave it to the physicist to find systems so that such results "scale".)
- Some H/ρ pairs may have analytic forms for their eigenpairs (but we really need cross-products to be "nice"). What could we do in that case?
- Short time optimization vs. steady-state optimization. Are we trying to optimize for sensing at some time t or at infinity?

Notation

If you have a problem

$$\text{minimize}\,\{f(x) : x \in \mathbb{R}^n\} \;\text{ when }\; f(x) = h(F(x)),$$

Notation

If you have a problem

$$\text{minimize}\{f(x) : x \in \mathbb{R}^n\} \text{ when } f(x) = h(F(x)),$$

- $h\colon \mathbb{R}^p \to \mathbb{R}$ is known
  - Smooth or nonsmooth
  - Convex or nonconvex
  - Has known derivative or subdifferential
  - Relatively cheap to evaluate

Notation

If you have a problem

$$\text{minimize}\,\{f(x) : x \in \mathbb{R}^n\} \text{ when } f(x) = h(F(x)),$$

- $h \colon \mathbb{R}^p \to \mathbb{R}$ is known
  - Smooth or nonsmooth
  - Convex or nonconvex
  - Has known derivative or subdifferential
  - Relatively cheap to evaluate
- $F \colon \mathbb{R}^n \to \mathbb{R}^p$ is relatively unknown
  - Based on a simulation
  - Relatively expensive to evaluate
  - Stochastic

Use knowledge of h to use fewer calls to F.

Contact

Thanks for listening! Questions?

jmlarson@anl.gov

This work was supported by the U.S. Department of Energy, Office of Science, Office of Advanced Scientific Computing Research, Accelerated Research for Quantum Computing program